

PAPER DETAILS

TITLE: DOĞAL DİL İŞLEMEDE DERİN ÖĞRENME UYGULAMALARI ÜZERİNE BİR LİTERATÜR
ÇALIŞMASI

AUTHORS: Dogan KÜÇÜK,Nursal ARICI

PAGES: 76-86

ORIGINAL PDF URL: <https://dergipark.org.tr/tr/download/article-file/610451>

DOĞAL DİL İŞLEMEDE DERİN ÖĞRENME UYGULAMALARI ÜZERİNE BİR LİTERATÜR ÇALIŞMASI

A LITERATURE STUDY ON DEEP LEARNING APPLICATIONS IN NATURAL LANGUAGE PROCESSING

Doğan KÜÇÜK*
Nursal ARICI**

Öz

Derin öğrenme, yapay zekâ ve makine öğrenmesi alanlarının önemli ve güncel bir konusu haline gelmiştir. Özellikle son yıllarda, farklı derin öğrenme yöntemleri öneren çalışmaların ve mevcut yöntemleri değişik problemler üzerinde uygulayan çalışmaların sayıları hızla artmaktadır. Doğal dil işleminin çeşitli alt alanlarında da bu yöntemler yaygın olarak kullanılmış ve halen kullanılmaktadır. Bu derleme çalışmasında, ilk olarak derin öğrenme yöntemlerinin bir sınıflandırması sunulmuş, ardından da doğal dil işleme problemlerine derin öğrenme yaklaşımlarının sunulduğu önemli çalışmalar incelenmiştir. Derin öğrenme ve doğal dil işleme problemlerinin çözümü amacıyla derin öğrenme konularıyla ilgili hem teorik çalışmaların hem de pratik uygulamalar içeren çalışmaların sayısının ve yaygınlığının daha da artacağı öngörülmektedir. Bu nedenle çalışmamızın; doğal dil işleme alanında derin öğrenme uygulamaları konusunda önemli bir Türkçe kaynak olacağı düşünülmektedir.

Anahtar Kelimeler: Doğal Dil İşleme, Metin Madenciliği, Derin Öğrenme, Makine Öğrenmesi.

Abstract

Deep learning is an important and recent topic of artificial intelligence and machine learning areas. Especially in recent years, the number of studies proposing different deep learning methods and applying these methods on different problems is increasing. These methods have also been used at various subareas of natural language processing extensively, and are still being used. In this survey paper, firstly, classification of deep learning techniques is presented and then important studies about deep learning approaches for natural language processing problems are discussed. It is expected that the number and prevalence of both theoretical studies and studies with practical applications on deep learning and on deep learning solutions to natural language processing problems are going to increase. Therefore it is considered that our study will be an important Turkish resource on the topic of deep learning applications for natural language processing.

Keywords: Natural Language Processing, Text Mining, Deep Learning, Machine Learning.

* Gazi Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği Bölümü, dogan.kucuk@gazi.edu.tr
ORCID: 0000-0001-5265-3263

** Doç. Dr. Gazi Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği Bölümü, nursal@gazi.edu.tr
ORCID: 0000-0002-4505-1341

1. GİRİŞ

Derin öğrenme, günümüzde oldukça popüler olan yapay zekâ ve makine öğrenmesi alanlarının önemli güncel bir alt alanıdır. Derin öğrenme metotları, genel olarak yapay sinir ağları (YSA - artificial neural networks) çalışmaları üzerine geliştirilmişlerdir. Ancak bu çalışmalardan farklı olarak daha fazla sayıda gizli nöron ve katman üzerine kuruludurlar (Ravi vd., 2017).

Derin öğrenme, eldeki verinin birden çok soyutlama seviyesinde temsil edilebilmesi için birden çok işlem katmanı bulunan hesaplama modellerini kapsamaktadır (Bengio, 2009; Lecun vd., 2015). Derin ağlar olarak da bilinen derin öğrenme yöntemlerinde, verinin temsili için üst üste olacak şekilde farklı katmanlar mevcuttur. Derin öğrenme yöntemleri ham verilerden etkin bir üst seviye soyutlaması yapmakla, otomatik öznitelik kümeleri oluşturabilmekte, böylelikle normalde çoğunlukla insanlar tarafından belirlenen özniteliklerin otomatik olarak çıkarılıp kullanılması sağlanmaktadır (Ravi vd., 2017).

Bazı derin öğrenme algoritmalarının öğrenme aşaması oldukça uzun sürebilmektedir ve farklı çalışmalarda, derin sinir ağlarının öğrenme süreçlerini kısaltmak için yarı-denetimli öğrenme yaklaşımları önerilmiştir (Lee, 2013). Bazı çalışmalarda ise destek vektör makinaları (support vector machines) gibi derin olmayan ama yaygın kullanılan makine öğrenmesi yöntemlerine derin öğrenmenin yeteneklerini kazandırma yönünde yaklaşımlar sunulmuştur (Cho ve Saul, 2009).

Derin öğrenme yöntemleri; video, ses, metin gibi birçok veri türünün işlenmesinde oldukça başarılı sonuçlar vermiştir (LeCun vd., 2015). Çoğunlukla bir yöntem, doğal dil işleme gibi metin verisi işleme konularında başarılı olabilirken (Collobert ve Weston, 2008), farklı bir yöntem video ve ses verilerinin işlenmesinde daha başarılı sonuçlar verebilmektedir. Bazı çalışmalarda ise, önerilen derin öğrenme yaklaşımları çok modlu (metin, görüntü ve ses gibi farklı modları aynı anda) öğrenmede başarıyla kullanılabilmektedir (Ngiam ve vd., 2011).

Derin öğrenme yöntemlerinin daha ayrıntılı olarak verilebilecek uygulama alanları arasında aşağıdakiler yer almaktadır (Deng ve Yu, 2014):

1. Dil modelleme ve doğal dil işleme (language modeling and natural language processing)
2. Konuşma ve ses işleme (speech and audio processing)
3. Bilgi erişimi (information retrieval)
4. Nesne tanıma ve bilgisayarlı görü (object recognition and computer vision)
5. Çok modlu ve çok görevli öğrenme (multimodal and multitask learning)

Yukarıdaki listede ilk madde olarak yer alan doğal dil işleme konusu, yapay zekâ alanı içerisinde değerlendirilmektedir ve metin madenciliği olarak da adlandırılmaktadır. Doğal dil işleme, doğal dilde oluşturulmuş metinler üzerinde otomatik olarak gerçekleştirilen metin ayrıştırma, metin sınıflandırma, bilgi çıkarımı, duygu analizi gibi birçok önemli konuyu bünyesinde barındırmaktadır ve uygulama alanları da gün geçtikçe artmaktadır.

Bu çalışmamızda; doğal dil işleme problemlerinin çözümü için derin öğrenme yöntemlerini kullanan çalışmalar incelenmiş olup Türkçe bir kaynak olarak araştırmacıların dikkatine sunulmuştur. Çalışmanın kalan bölümleri şu şekilde ilerlemektedir: 2. bölümde derin öğrenme yöntemleri hakkında genel bilgiler verilmiş; 3. bölümde doğal dil işleme alanında yapılmış derin öğrenme çalışmaları incelenmiş, 4. bölümde ise makale özetlenmiş ve sonuçlar verilmiştir.

2. DERİN ÖĞRENME YÖNTEMLERİ

Bu bölümde aşağıda listelenen altı derin öğrenme yöntemi (Ravi vd., 2017) incelenmiş ve bunların uygulama alanları hakkında bilgiler verilmiştir:

1. Derin Sinir Ağları (Deep Neural Networks - DNN)
2. Derin Oto-kodlayıcılar (Deep Autoencoders)
3. Derin İnanç Ağları (Deep Belief Networks - DBN)
4. Derin Boltzmann Makinesi (Deep Boltzmann Machine - DBM)
5. Yinelene Sinir Ağları (Recurrent Neural Networks - RNN)
6. Evrişimsel Sinir Ağları (Convolutional Neural Networks - CNN)

Aşağıda Tablo 1’de bu yöntemlerin niteliksel bir karşılaştırılması (Ravi vd., 2017) çalışmasından yararlanılarak sunulmuştur. Çok katmanlı yapılarından dolayı söz konusu derin öğrenme algoritmalarının tümü için öğrenme süreci hesaplama anlamında maliyetlidir. Algoritmaların farklı alanlarda uygulanmaları sırasında, uygulamanın gereksinimleri, yapılan parametre ayarlamaları ve kullanılan veri kümesinin boyutuna göre performansları farklılık gösterebilmektedir. Dolayısıyla, yöntemlerin karşılaştırması Tablo 1’de sadece niteliksel anlamda sunulmuştur.

Tablo 1. Tanıtılan Derin Öğrenme Yöntemlerinin Niteliksel Karşılaştırılması

<i>Derin Öğrenme Yöntemi</i>	<i>Özellikleri</i>
Derin Sinir Ağları	Birçok alanda başarıyla uygulanmaktadır. Ancak öğrenme süreci uygulama alanına bağlı olarak oldukça yavaş olabilmektedir.
Derin Oto-kodlayıcılar	Öğrenme için işaretli veri gerektirmeyen denetimsiz bir derin öğrenme yöntemidir.
Derin İnanç Ağları	Denetimli ve denetimsiz şekilde kullanılabilir. Ancak eğitim süreci oldukça yavaş olabilmektedir.
Derin Boltzmann Makinesi	Derin İnanç Ağları’ndan farkı, sadece en üst iki katman dışında tüm katmanlar arasında da yönsüz bağlantıların olmasıdır. Bu nedenle de Derin İnanç Ağları’na göre hesaplama anlamında daha maliyetlidir.
Yinelene Sinir Ağları	Sıralı desenleri belirlemede başarılıdır. Bu nedenle özellikle Long Short-Term Memory (LSTM) gibi alt türleri doğal dil işleme gibi alanlarda başarılı sonuçlar vermektedir.
Evrişimsel Sinir Ağları	Genellikle bilgisayarlı görü alanında ve görsel veriler üzerinde kullanılmakta ve büyük boyutta işaretli veri gerektirmesine rağmen başarılı sonuçlar vermektedir. Doğal dil işleme alanındaki problemlere de başarıyla uygulanmıştır.

2.1. Derin Sinir Ağları

Literatürde, sinir ağları konusundaki ilk algoritmalarından biri perseptron algoritmasıdır (Rosenblatt, 1958). Bu ağda, bir girdi katmanı mevcuttur ve direk olarak çıktıya bağlıdır. Bu algoritma ile lineer olarak ayrılabilen desenlerin ayrıştırılması amacıyla sınıflandırıcılar geliştirilebilir (Ravi vd., 2017). Daha karmaşık problemler için bu algoritmaya birden fazla gizli katman eklenmiş ve delta kuralı adı verilen öğrenme yöntemiyle her katmanın ağırlığı ayarlanabilmiştir.

Bu tipteki sinir ağlarına daha fazla sayıda gizli katman (ikiden fazla) eklenmesiyle doğrusal olmayan karmaşık ilişkilerin de tespit edilebilmesi sağlanmıştır ve elde edilen bu sinir ağlarına derin sinir ağları (DSA) adı verilmiştir. DSA'lar hem (işaretli veriler varsa) denetimli, hem de kümeleme gibi denetimsiz öğrenme problemleri için kullanılabilir. DSA'lar yaygın olarak sınıflandırma ve regresyon amacıyla kullanılmakta ve başarılı sonuçlar elde edilmesini sağlamaktadırlar. Ancak öğrenme süreçleri oldukça yavaş olabilmektedir (Ravi vd., 2017).

2.2. Derin Oto-Kodlayıcılar

Oto-kodlayıcılar, problem çözümü için gerekli öznelik kümesinin veriden otomatik olarak çıkarılması amacıyla ortaya atılmış sinir ağlarıdır. Bir oto-kodlayıcı, girdi vektörüne bir sınıf etiketi vermek yerine onu yeniden oluşturmak için öğretilmektedir (Ravi vd., 2017).

Derin öğrenmenin konusu olan derin oto-kodlayıcılar ise, çok boyutlu veriyi temsil etmek için birden fazla oto-kodlayıcının birbiri üstüne kümelenmesiyle oluşan mimarilerdir (Hinton ve Salakhutdinov, 2006). Literatürde, derin oto-kodlayıcıların birçok farklı türü önerilmiştir.

Derin oto-kodlayıcıların amacı, öznelik kümesini otomatik çıkarmak veya veri boyutu sayısını azaltmaktır (Hinton ve Salakhutdinov, 2006). Bu yöntemde öğrenme için işaretli veri kümesine ihtiyaç yoktur, yani denetimsiz bir öğrenme yöntemidir. Ancak yöntem uygun ağırlıkları bulabilmek için bir ön-öğrenme aşamasına ihtiyaç duymaktadır. Bu ön-öğrenme aşamasında son konfigürasyona uygun yaklaşık ağırlıklar yönteme sağlanmaktadır (Ravi vd., 2017).

2.3. Derin İnanç Ağları

Derin inanç ağları (Hinton vd., 2006) ve bir sonraki bölümde tanıtılacak olan derin Boltzmann makineleri, sınırlandırılmış Boltzmann makinesi adı verilen bir algoritmaya (Hinton ve Sejnowski, 1986) dayanmaktadır. SBM algoritması stokastik bir sinir ağı olarak tanımlanmakta ve bu ağlarda Gaussian gibi belirli bir dağılıma sahip stokastik birimler kullanılmaktadır. Öğrenme sürecinde ise Gibbs örnekleme gibi ağırlıkları adım adım ayarlayan yöntemler kullanılmaktadır.

Derin inanç ağları birden fazla SBM'nin bileşkesi olarak algılanabilir (Hinton vd., 2006). Her bir SBM'nin gizli katmanı, bir sonraki SBM'nin görünür katmanına bağlanmıştır ve en üst seviyede yönsüz bağlantılar vardır. Derin inanç ağları hem denetimli hem de denetimsiz öğrenme amacıyla kullanılabilir. Bu ağı başlatmak için katman-katman bir açgözlü algoritma ile öğrenme gerçekleştirilmektedir (Ravi vd., 2017).

2.4. Derin Boltzmann Makinesi

Boltzmann makinelerine dayalı bir diğer derin öğrenme yöntemi de derin Boltzmann makineleridir (Salakhutdinov ve Larochelle, 2010). Derin inanç ağlarından farkı; derin Boltzmann makinelerinde ağın tüm katmanları arasında yönsüz bağlantıların olmasıdır. Bir diğer fark ise derin Boltzmann makinelerinde zaman karmaşıklığının daha fazla olması bu nedenle de büyük veri kümelerinde öğrenme sürecinin yavaş olmasıdır (Ravi vd., 2017).

2.5. Yinelenen Sinir Ağları

Yinelenen sinir ağları (Williams ve Zipser, 1989), veri akışlarını (stream of data) analiz edebilen gizli katmanlara sahip sinir ağlarıdır ve çıktının bir önceki hesaplamalara bağlı olduğu problemlerin çözümü için çok uygundur (Ravi vd., 2017). Yinelenen sinir ağları bu nedenle özellikle doğal dil işleminin değişik problemlerinin çözümünde oldukça başarılı olmuştur (Ravi vd., 2017). Yinelenen sinir ağlarında öğrenme sırasında ortaya çıkan bazı problemler nedeniyle, bu ağların uzun kısa-dönem bellek (Long Short-Term Memory - LSTM) (Hochreiter ve Schmidhuber, 1997) gibi farklı sürümleri literatüre girmiştir.

2.6. Evrişimsel Sinir Ağları

Evrişimsel sinir ağları, çok boyutlu girdiler için ve özellikle iki boyutlu görsel veriler için önerilmiş bir derin öğrenme yöntemidir (LeCun vd., 1998). Evrişimsel sinir ağları diğer sinir ağlarına göre daha az sayıda nöron bağlantısına sahiptir ve birçok farklı sürümü literatürde mevcuttur. Bu sinir ağları öğrenme aşamasında oldukça büyük boyutta işaretli veriye ihtiyaç duymaktadır (Ravi vd., 2017). Bu derin öğrenme yöntemi doğal dil işleme alanındaki çeşitli problemler üzerinde uygulanmış ve başarılı sonuçlar elde edilmiştir.

3. DOĞAL DİL İŞLEME ALANINDAKİ DERİN ÖĞRENME ÇALIŞMALARI

Derin öğrenmenin bir önceki bölümde anlatılan farklı yöntemlerinden özellikle yinelenen sinir ağları ve evrişimsel sinir ağlarının doğal dil işleme problemlerini çözmede de başarılı sonuçlar verdiği görülmüştür. Örneğin; evrişimsel sinir ağlarının doğal dil işlemenin çok farklı problemlerini çözmede kullanılabileceği, çok görevli öğrenme için birleşik bir mimarinin tanıtıldığı bir çalışmada gösterilmiştir (Collobert ve Weston, 2008). Bahsi geçen bu iki yöntemle birlikte; diğer derin öğrenme yöntemlerinin de doğal dil işlemenin çeşitli problemlerinde kullanılabileceği ve bu yöntemlere dayalı başarılı sistemlerin geliştirilebileceği, literatürdeki farklı çalışmalarda ifade edilmiştir (Socher vd., 2012).

Bu bölümün aşağıdaki alt bölümlerinde de doğal dil işleme alanındaki örnek problemler ve bu problemler için derin öğrenme yaklaşımlarının tanıtıldığı çalışmalara yer verilmiştir. Ayrıca, Tablo 2’de bu alt bölümlerde açıklanan problemler ve çözümünde kullanılan derin öğrenme yöntemleri özet şeklinde sunulmuştur.

Tablo 2. Tanıtılan Doğal Dil İşleme Problemleri ve Kullanılan Derin Öğrenme Yöntemleri

<i>Doğal Dil İşleme Problemi</i>	<i>Kullanılan Derin Öğrenme Yöntemleri</i>
Metin Sınıflandırma	• Evrişimsel Sinir Ağları • Yinelemeli Evrişimsel Sinir Ağları • LSTM ve Evrişimsel Sinir Ağları
Metin Ayrıştırma	• Evrişimsel Sinir Ağları
Duygu Analizi	• Derin Oto-Kodlayıcılar • Evrişimsel Sinir Ağları
Bilgi Çıkarımı	• Derin Sinir Ağları
Varlık İsmi Tanıma	• LSTM ve Evrişimsel Sinir Ağları
Zamansal İlişki Çıkarımı	• Evrişimsel Sinir Ağları
Olay Çıkarımı	• Evrişimsel Sinir Ağları
Sözcük Türü Etiketleme	• Derin Sinir Ağları • LSTM
Metin Sıralama	• Evrişimsel Sinir Ağları
Otomatik Harf Çevirisi	• Derin İnanç Ağları
Otomatik Soru Cevaplama	• Evrişimsel Sinir Ağları • LSTM

3.1. Metin Sınıflandırma

Doğal dil işlemenin uzun zamandır çalışılan ve önemli uygulama alanları olan konularından bir tanesi metin sınıflandırmadır. Derin öğrenme yöntemleri, metin sınıflandırma amacıyla da kullanılmaktadır ve bu konuda çeşitli güncel çalışma mevcuttur.

Örnek bir çalışmada, değişik veri kümeleri üzerinde cümle sınıflandırma problemi (olumlu/olumsuz müşteri değerlendirmeleri gibi) için basit tek seviyeli bir evrimsel sinir ağı yaklaşımı ile oldukça başarılı sonuçlar alınmıştır (Kim, 2014).

Bir diğer örnek çalışmada, metin sınıflandırma amacıyla bir yinelemeli evrimsel sinir ağı altyapısı önerilmiştir. Önerilen yaklaşımda; bağlamsal bilgi için yinelemeli yapı, metnin temsili içinse evrimsel sinir ağı kullanılmıştır (Lai vd., 2015).

Yine metin sınıflandırma konusundaki bir başka çalışmada, karakter seviyesinde evrimsel ağlar kullanılmıştır. Çalışmada, karakter seviyesindeki evrimsel sinir ağlarının metin sınıflandırma etkili bir yöntem olduğu ve bu yöntemin benzer problemlerin çözümünde de kullanılabileceği ifade edilmiştir (Zhang vd., 2015).

Başka bir çalışmada ise sıralı kısa metinlerin sınıflandırılması için yinelemeli sinir ağının bir türü olan LSTM ve evrimsel sinir ağı yöntemleri kullanılarak çeşitli veri kümelerinde test edilerek başarılı sonuçlar elde edilmiştir (Lee ve Dernoncourt, 2016).

3.2. Metin Ayırıştırma

Otomatik metin ayırıştırma, doğal dil işlemenin çok uzun yıllardır çalışılan ve çözülmesi amacıyla birçok farklı yöntemin literatüre kazandırıldığı bir alanıdır. Metin ayırıştırma verilen bir metnin gramatik yapısının ortaya çıkarılması amaçlanmaktadır.

Bu konuda gerçekleştirilen bir çalışmada, metin ayırıştırma için evrimsel sinir ağlarına dayalı ve hızlı bir yöntem önerilmiştir. Bu yöntemde evrimsel sinir ağları, grafiklerde yapılandırılmış etiket çıkarımı ile birleştirilmiş ve böylelikle literatürde daha önce mevcut olan grafik dönüştürücü ağ yöntemi kullanılmıştır. Önerilen yöntemin geliştirmeye açık yönleri bulunsun da, metin ayırıştırma için oldukça etkin bir yöntem olduğu ifade edilmiştir (Collobert, 2011).

3.3. Duygu Analizi

Duygu analizi, verilen bir metindeki duygu, fikir ve özneliğin otomatik olarak belirlenmesi olarak tanımlanmaktadır (Pang ve Lee, 2008). Yaygın olarak, duygu analizi çalışmalarında verilen metin kullanılarak bu metnin olumlu, olumsuz veya tarafsız şeklinde sınıflandırılması amaçlanmaktadır. Duygu analizi, doğal dil işlemenin uzun süredir çalışılan ve halen de yaygın olarak çalışılmaya devam edilen güncel araştırma konularından biridir. Bu alandaki güncel çalışmalardan bazılarında da derin öğrenme yöntemleri kullanılmaktadır.

Bu konudaki bir çalışmada, duygu analizi yaklaşımlarını alandan bağımsız hale getirebilmek ve böylelikle yeni metin türlerine de adapte edebilmek için oto-kodlayıcı tabanlı bir derin öğrenme yöntemi kullanılmıştır. Kullanılan oto-kodlayıcı yaklaşımıyla denetimsiz şekilde farklı metin türleri için gerekli duygu analizi öznelikleri otomatik olarak belirlenmektedir ve oldukça başarılı sonuçlar elde edilmiştir (Glorot vd., 2011). Bir diğer ilgili çalışmada yine derin oto-kodlayıcılar yarı-denetimli (semi-supervised) bir yaklaşımla eğitilerek duygu analizinde başarılı sonuçlar elde etmişlerdir (Zhai ve Zhongfei, 2016).

Bir diğer örnek çalışmada, evrimsel sinir ağları kullanılarak Twitter üzerinde duygu analizi gerçekleştiren bir yaklaşım anlatılmıştır. Söz konusu çalışmada üç aşamalı ve her aşamada farklı öğrenme yöntemlerinin kullanıldığı bir yaklaşım benimsenmiş, tek seviyeli basit bir evrimsel sinir

ağı kullanılmıştır (Severyn ve Moschitti, 2015a). Çalışma tweet veri kümeleri üzerinde denenmiş ve başarılı sonuçlar elde edilmiştir.

İlgili bir diğer çalışmada, kısa metinlerde (film yorumları ve tweet'ler) duygu analizi için yine evrişimsel sinir ağı yaklaşımı benimsenmiştir (dos Santos ve Gatti, 2015). Bu yaklaşımda iki evrişimsel katman vardır. Bu katmanlarda sözcüklerden karakter seviyesinde ve cümlelerden cümle seviyesindeki öznitelikler çıkarılmaktadır. Çalışmada önerilen yaklaşım film yorumları ve tweet'lerden oluşan iki ayrı veri kümesi üzerinde denenmiş ve başarılı sonuçlar elde edilmiştir (dos Santos ve Gatti, 2015).

3.4. Bilgi Çıkarımı

Bilgi çıkarımı, doğal dil işlemenin önemli araştırma alanlarından bir tanesidir. Bilgi çıkarımı ile doğal dildeki metinlerden; varlıklar, olaylar, tarih ve zaman ifadeleri, kavramlar gibi birçok kayda değer bilgi parçası otomatik şekilde çıkarılmaktadır.

Bilgi çıkarımı için derin öğrenme yöntemi kullanılan bir çalışmada, çıkarılan bilginin türünden bağımsız genel bir yaklaşım önerilmiştir (Qi vd., 2014). Söz konusu çalışmada, bilgi çıkarımı için derin sinir ağları kullanılmış ve özniteliklerin elle oluşturulması yerine özniteliklerin öğrenilmesi sağlanmaktadır. Çalışma, farklı bilgi çıkarımı veri kümelerinde denenmiş ve başarılı sonuçlar elde etmiştir (Qi vd., 2014).

3.5. Varlık İsmi Tanıma

Varlık ismi tanıma, doğal dil işlemenin bilgi çıkarımı alanının bir alt problemidir ve uzun süredir çalışılan bir araştırma konusudur. Varlık ismi tanıma sistemleri, genel olarak, verilen bir metindeki kişi, yer ve kurum gibi varlıkların isimlerini otomatik olarak çıkarıp sınıflandırır.

Derin öğrenme yöntemlerinin varlık ismi tanıma da başarılı sonuçlar elde ettiği görülmüştür. Bu konudaki bir çalışmada, varlık ismi tanıma için hibrit ve çift yönlü bir yaklaşım önerilerek hem yinelenen sinir ağlarının bir algoritması olan LSTM, hem de evrişimsel sinir ağları kullanılmıştır. Kullanılan bu hibrit derin öğrenme yaklaşımıyla kelime seviyesindeki ve karakter seviyesindeki öznitelikler otomatik olarak belirlenmektedir. Söz konusu yaklaşım, varlık ismi tanıma alanındaki ortak veri kümeleri üzerinde denenmiş ve oldukça başarılı sonuçlar elde edilmiştir (Chiu ve Nichols, 2015).

3.6. Zamansal İlişki Çıkarımı

Zamansal ilişkilerin çıkarımı da bilgi çıkarımı konusunun bir alt konusudur ve dolayısıyla doğal dil işlemenin önemli bir araştırma alanıdır. Zamansal ilişkiler arasında şu ve benzeri ifadeler yer almaktadır: “*önce, sonra, hemen önce, hemen sonra, sırasında (esnasında)*”.

Bu konuyla ilgili yapılmış bir çalışmada, zamansal ilişki ifadelerinin çıkarılarak sınıflandırılması amacıyla evrişimsel sinir ağları kullanılmıştır. Çalışmada sözlük seviyesindeki ve cümle seviyesindeki öznitelikler kullanılmıştır. Ancak, çalışmada literatürdeki benzer çalışmalardan daha iyi sonuçlar elde edilememiştir ve çeşitli ön işleme (preprocessing) yöntemleri kullanılarak sonuçların iyileştirilebileceği ifade edilmiştir (Do ve Jeong, 2016).

3.7. Olay Çıkarımı

Doğal dilde oluşturulmuş metinlerde bahsedilen olayların otomatik çıkarımı da bilgi çıkarımı konusunun bir alt alanıdır ve dolayısıyla doğal dil işlemenin bir problemidir.

Bu alanda yapılmış güncel bir çalışmada, otomatik olay çıkarımı için cümle seviyesindeki özniteliklerin çıkarılması için evrişimsel sinir ağlarının kullanıldığı bir yaklaşım tanıtılmıştır. Bu çalışmada ayrıca, kullanılan kelime seviyesindeki öznitelikler içinse farklı bir kelime temsili

yaklaşımı benimsenmiştir. Olay çıkarımı için derin öğrenme yöntemi kullanan bu çalışmanın olay çıkarımı deneylerinde başarılı sonuçlar verdiği gösterilmiştir (Chen vd., 2015).

3.8. Sözcük Türü Etiketleme

Sözcük türü etiketleme, doğal dilde verilen bir metindeki kelimelerin türlerinin sınıflandırılması problemidir. Sözcük türü sınıfları isim, fiil, edat, zamir gibi türlerdir. Sözcük türü etiketleme doğal dil işlemenin çok uzun süredir çalışılan bir konusudur ve ayrıca daha kapsamlı problemlerin çözümünde de söz konusu sözcük türleri kullanılabilir.

Bu konuda yapılmış bir çalışmada, derin sinir ağları kullanılarak sözcük türü belirleme amacıyla sözcük seviyesinde ve karakter seviyesindeki temsillerin öğrenilmesi sağlanmıştır. Söz konusu sinir ağına, öznitelik çıkarımı için evrimsel bir katman kullanılmaktadır. İngilizce ve Portekizce üzerinde yapılan deneylerde oldukça başarılı sonuçlar elde edilmiştir (dos Santos ve Zadrozny, 2014).

Sözcük türü etiketleme konusunda yapılmış dilden bağımsız bir çalışmada ise çift yönlü LSTM ağları kullanılmıştır. 22 farklı dil üzerinde yapılan testler sonucunda söz konusu derin öğrenme metoduyla başarılı sonuçlar elde edildiği gözlenmiştir (Plank vd., 2016).

3.9. Metin Sıralama

Metin sıralama, bilgi erişiminin kapsamında bulunan önemli bir araştırma konusudur. Bu sıralama işlemi için de derin öğrenme yöntemini kullanan çalışmalar mevcuttur.

Bu yöndeki bir çalışmada, evrimsel derin öğrenme yöntemi kullanılarak kısa metinlerin sıralanması konusu incelenmiştir. Bu sıralama işlemindeki ana konu, metinlerin sıralanması için kullanılacak benzerlik işlevinin öğrenilmesidir. Söz konusu çalışmada kullanılan derin öğrenme yöntemiyle, elle öznitelik çıkarma işlemi veya dış kaynaklar kullanılmasına gerek kalmadığı ifade edilmiştir. Önerilen yaklaşım, kısa metinlerin sıralanması deneylerinde başarılı sonuçlara ulaşmıştır (Severyn ve Moschitti, 2015b).

3.10. Otomatik Harf Çevirisi

Otomatik çeviri alanında kullanılan otomatik harf çevirisi alanında da derin öğrenme yöntemlerinin kullanıldığı görülmektedir. Otomatik harf çevirisi ile kelimelerin (özel isimler gibi) bir alfabedeki harflerden diğer alfabedeki harflere çevrilmesi sağlanmaktadır.

Bu konuda yapılmış bir çalışmada; derin inanç ağları yöntemi, Arapça-İngilizce arası harf çevirisinde kullanılmıştır. Bu konuda; daha önceki çalışmalarda, sonlu-durum makinaları ve ifade-tabanlı yöntemler kullanılmıştır. Derin inanç ağları, önceki çalışmalardaki bu yöntemlerden daha başarılı sonuçlar vermese de mevcut diğer yöntemlerle beraber kullanılabileceği ifade edilmiştir (Deselaers vd., 2009).

3.11. Otomatik Soru Cevaplama

Otomatik soru cevaplama sistemleri, doğal dilde ifade edilmiş bir sorunun cevabını doğal dildeki büyük veri kümeleri içerisinde bularak yine doğal dilde otomatik olarak cevaplamayı amaçlayan sistemlerdir.

Derin öğrenme yöntemleri soru cevaplama sistemlerinde de kullanılmıştır. Örneğin, Yu vd. (2014) tarafından gerçekleştirilen çalışmada, soru cevaplama işleminin bir alt işlemi olan cevap cümlesi seçimi için evrimsel sinir ağları kullanılmış ve karşılaştırılan mevcut başarılı sistemlere yakın sonuçlar elde edilmiştir.

İlgili bir diğer çalışmada yine soru cevaplama probleminin çözümünde kullanılmak üzere cevap seçimi için çift yönlü LSTM algoritması ile evrimsel sinir ağlarını kullanılmıştır. Önerilen

birleşik yöntem, karşılaştırmada kullanılan diğer algoritmalara göre başarılı sonuçlar elde etmiştir(Tan vd., 2015).

4. SONUÇ

Derin öğrenme, son dönemde oldukça popüler olmuş önemli bir araştırma alanıdır. Sinir ağlarına dayalı olan derin öğrenme konusunda literatürde birçok farklı yöntem önerilmiştir ve bunların bazıları özellikle bazı araştırma alanlarında oldukça başarılı sonuçlar elde edilmesini sağlamıştır. Doğal dil işleme alanı da söz konusu önemli araştırma alanlarından biridir.

Bu çalışmamızda, doğal dil işleme için derin öğrenme yöntemleri konusunda yapılmış çalışmaların derlenmesi amaçlanmıştır. Çalışmamızda ilk olarak, derin öğrenme konusunda temel bilgiler ile literatürde mevcut bulunan altı önemli derin öğrenme yöntemi hakkında detaylı bilgiler ve bunların belli başlı uygulama alanları verilmiştir. Ardından da doğal dil işleme için derin öğrenme yöntemlerini tanıtan çalışmalar hakkında bilgiler sunulularak, yöntemlerin başarımları ve yetersiz kaldığı durumlar açıklanmıştır. Söz konusu derin öğrenme çalışmaları, ilgili doğal dil işleme problemlerine göre mümkün olduğunca sınıflandırılmış, ayrı alt başlıklar altında verilmiştir. Çalışmamızın, Türkçe doğal dil işleme alanında derin öğrenme yöntemlerini uygulayacak araştırmacılara Türkçe bir kaynak olarak ışık tutması ve yol gösterici olması beklenmektedir.

KAYNAKÇA

- Bengio, Y. (2009). Learning deep architectures for AI. Foundations and Trends in Machine Learning 2(1), 1-127.
- Chen, Y., Xu, L., Liu, K., Zeng, D., & Zhao, J. (2015). "Event extraction via dynamic multi-pooling convolutional neural networks". Annual Meeting of the Association for Computational Linguistics and International Joint Conference on Natural Language Processing, 167-176.
- Chiu, J. P., & Nichols, E. (2015). "Named entity recognition with bidirectional LSTM-CNNs". arXiv preprint arXiv:1511.08308.
- Cho, Y., & Saul, L. K. (2009). "Kernel methods for deep learning". Advances in Neural Information Processing Systems, 342-350.
- Collobert, R. (2011). "Deep learning for efficient discriminative parsing". International Conference on Artificial Intelligence and Statistics, 224-232.
- Collobert, R., & Weston, J. (2008). "A unified architecture for natural language processing: Deep neural networks with multitask learning". International Conference on Machine Learning (ICML), 160-167.
- Conneau, A., Schwenk, H., Barrault, L., & Lecun, Y. (2016). "Very deep convolutional networks for natural language processing". arXiv preprint arXiv:1606.01781.
- Deng, L., & Yu, D. (2014). "Deep learning: methods and applications". Foundations and Trends in Signal Processing, 7(3-4), 197-387.
- Deselaers, T., Hasan, S., Bender, O., & Ney, H. (2009). "A deep learning approach to machine transliteration". International Workshop on Statistical Machine Translation, 233-241.
- Do, H. W., & Jeong, Y. S. (2016). "Temporal relation classification with deep neural network". International Conference on Big Data and Smart Computing (BigComp), 454-457.

- Dos Santos, C. N., & Gatti, M. (2014). "Deep convolutional neural networks for sentiment analysis of short texts". *International Conference on Computational Linguistics (COLING)*, 69-78.
- Dos Santos, C. N., & Zadrozny, B. (2014). "Learning character-level representations for part-of-speech tagging". *International Conference on Machine Learning (ICML)*, 1818-1826.
- Glorot, X., Bordes, A., & Bengio, Y. (2011). "Domain adaptation for large-scale sentiment classification: A deep learning approach". *International Conference on Machine Learning (ICML)*, 513-520.
- Hinton, G. E., & Salakhutdinov, R. R. (2006). "Reducing the dimensionality of data with neural networks". *Science*, 313(5786), 504-507.
- Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). "A fast learning algorithm for deep belief nets". *Neural Computation*, 18(7), 1527-1554.
- Hinton, G. E., & Sejnowski, T. J. (1986). "Learning and relearning in boltzmann machines". *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, 1(282-317), 2.
- Hochreiter, S., & Schmidhuber, J. (1997). "Long short-term memory". *Neural Computation*, 9(8), 1735-1780.
- Kim, Y. (2014). "Convolutional neural networks for sentence classification". *arXiv preprint arXiv:1408.5882*.
- Lai, S., Xu, L., Liu, K., & Zhao, J. (2015). "Recurrent convolutional neural networks for text classification". *AAAI Conference on Artificial Intelligence*, 2267-2273.
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). "Deep learning". *Nature* 521(7553), 436.
- Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). "Gradient-based learning applied to document recognition". *Proceedings of the IEEE*, 86(11), 2278-2324.
- Lee, D. H. (2013). "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks". *Workshop on Challenges in Representation Learning, ICML (3)*, 2
- Lee, J. Y., & Deroncourt, F. (2016). "Sequential short-text classification with recurrent and convolutional neural networks". *arXiv preprint arXiv:1603.03827*.
- Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., & Ng, A. Y. (2011). "Multimodal deep learning". *International Conference on Machine Learning*, 689-696.
- Pang, B., & Lee, L. (2008). "Opinion mining and sentiment analysis". *Foundations and Trends in Information Retrieval*, 2(1-2), 1-135.
- Plank, B., Søgaard, A., & Goldberg, Y. (2016). "Multilingual part-of-speech tagging with bidirectional long short-term memory models and auxiliary loss". *arXiv preprint arXiv:1604.05529*.
- Qi, Y., Das, S. G., Collobert, R., & Weston, J. (2014). "Deep learning for character-based information extraction". *European Conference on Information Retrieval*, 668-674.
- Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., & Yang, G. Z. (2017). "Deep learning for health informatics". *IEEE journal of Biomedical and Health Informatics* 21(1), 4-21.
- Rosenblatt, F. (1958). "The perceptron: a probabilistic model for information storage and organization in the brain". *Psychological Review*, 65(6), 386.

- Salakhutdinov, R., & Larochelle, H. (2010). "Efficient learning of deep Boltzmann machines". International Conference on Artificial Intelligence and Statistics.
- Severyn, A., & Moschitti, A. (2015) a. "Twitter sentiment analysis with deep convolutional neural networks". International ACM SIGIR Conference on Research and Development in Information Retrieval, 959-962.
- Severyn, A., & Moschitti, A. (2015) b. "Learning to rank short text pairs with convolutional deep neural networks". International ACM SIGIR Conference on Research and Development in Information Retrieval, 373-382.
- Socher, R., Bengio, Y., & Manning, C. D. (2012). "Deep learning for NLP (without magic)". Tutorial Abstracts of ACL 2012, 5.
- Tan, M., Santos, C. D., Xiang, B., & Zhou, B. (2015). "LSTM-based deep learning models for non-factoid answer selection". arXiv preprint arXiv:1511.04108.
- Williams, R. J., & Zipser, D. (1989). "A learning algorithm for continually running fully recurrent neural networks". Neural Computation, 1(2), 270-280.
- Yu, L., Hermann, K. M., Blunsom, P., & Pulman, S. (2014). "Deep learning for answer sentence selection". arXiv preprint arXiv:1412.1632.
- Zhai, S., & Zhang M. Z. (2016). "Semisupervised autoencoder for sentiment analysis". AAAI Conference on Artificial Intelligence, 1394-1400.
- Zhang, X., Zhao, J., & Lecun, Y. (2015). "Character-level convolutional networks for text classification". Advances in Neural Information Processing Systems, 649-657.